Volume Title: Annals of Economic and Social Measurement, Volume 5, number 3

Volume Author/Editor: NBER

Volume Publisher:

Volume URL: http://www.nber.org/books/aesm76-3

Publication Date: July 1976

Chapter Title: First Order Dual Control

Chapter Author: Alfred L. Norman

Chapter URL: http://www.nber.org/chapters/c10482

Chapter pages in book: (p. 311 - 321)

Annals of Economic and Social Measurement, 5/3, 1976

FIRST ORDER DUAL CONTROL[1]

BY ALFRED L. NORMAN

For large econometric models, computational simplicity is a desirable property of active learning strategies. This paper presents and evaluates one such strategy, first order dual control, DUAL1. In the development of DUAL1 the unknown parameters are treated directly without augmentation to the states. To calculate the current period control requires only one calculation of the Ricatti system. The Monte Carlo comparisons with two passive learning strategies, heuristic certainty equivalence, HCE and open loop mean variance, OLMV, indicate the relative performance of the HCE and OLMV strategies is problem specific and there exist problems where parameter estimation error can lead to poorer performance for the DUAL1 strategy than the OLMV strategy.

I. INTRODUCTION

Consider the following stochastic control problem:

Determine

$$J^*(X_{j-1}) = \min_{U_j U_{j+1} \ldots U_T} J(X_{j-1})$$

where

$$(1.1)\qquad J = E\left\{\sum_{t=j}^{T} \left[\tfrac{1}{2}(X_t-\alpha_t)' W_1 (X_t-\alpha_t) + \tfrac{1}{2}(U_t-\beta_t)' W_2 (U_t-\beta_t) \middle| \mathscr{P}_{t-1}\right]\right\}$$

subject to

$$(1.2)\qquad X_t = AX_{t-1} + BU_t + CZ_t + \epsilon_t$$

with the following observation pattern: $\mathscr{P}_{t-1}$:X_t, $k=-N+1, N+2, \ldots t-1$ is observed without error prior to executing U_t, and where

- X_t is an n-vector of state variables,
- U_t is an m-vector of control variables,
- Z_t is an r-vector of exogenous variables which are assumed known throughout the planning horizon,
- ε_t is an n-vector of disturbances with the following characteristics $E\varepsilon_t = 0$, $E\varepsilon_t\varepsilon'_t = \Omega$, ε_t, and ε_s are statistically independent.
- A, B, C, are $n\times n$, $n\times m$, and $n\times r$ matrices respectively. These matrices which contain unknown constant elements can arise directly from a model specified as a reduced form (1.2) a model specified as a structural form

$$(1.3)\qquad A_0X_t = A_sX_{t-1} + B_sU_t + C_sZ_t + \varepsilon_t$$

in which case, assuming A_0^{-1} exists, $A = A_0^{-1}A_s$, $B = A_0^{-1}B_s$, $C = A_0^{-1}C_s$.

- W_1 and W_2 are symmetric weighing matrices and $[B'W_1B + W_2]$ is positive definite.

[1] Research supported by NSF Grant Soc. 72-05254.

Prior to the control experiment the system (1.2) has been observed for N periods under a regime of nonoptimal control.

For large conometric models there is a need for computationally simple active learning strategies for non-Bayesian estimation. The purpose of this paper is to present and evaluate one such estimation and control strategy, first order dual control, DUAL1.

For the stochastic control problem under consideration the optimal stochastic control law is not computable. To formulate an estimation and control strategy requires replacing the unknown parameters with proxy variables. Active learning strategies are based on replacing the unknown parameters with random variables whose means equal the parameter estimates and whose covariances are based on the actual data plus anticipated path. The Bayesian dual control strategy [1], [9], could be adapted to non-Bayesian estimation; however, this approach has a major disadvantage for large econometric models. For unknown parameters Tse and Bar-Shalom augment the state vector. As the computation of the Ricatti matrices is cubic in the number of states, [5], augmenting the state will incur large computational costs for a large econometric model. The first order dual control strategy, which is derived in Appendix 1, approaches the estimation and control problem without augmenting the unknown parameters to the states. As shown in Appendix 1, the Ricatti matrices for the linear and quadratic term are equivalent for the deterministic and perturbation control. This implies that the Ricatti matrices need be computed only once to compute the current period control.

Two Monte Carlo experiments were designed to test the performances of DUAL1, with two passive learning strategies, heuristic certainty equivalence, HCE, and open loop mean variance, OLMV. In the HCE strategy the unknown parameters are replaced with the estimates, which are updated with each new observation. In the literature [2], [4], [7], HCE is generally known as certainty equivalence, CE. The adjective heuristic is added to emphasize the fact this strategy is generally not optimal. HCE is also known as linear decision rule, LDR, [6], and also forced separation [3]. In the OLMV strategy the unknown parameters are replaced by statistically independent random variables whose means equal the parameter estimates and whose covariances equal the estimate covariances. The means and covariances, which are updated with each observation, are assumed fixed over the planning horizon. OLMV is known as uncertainty adverse [4], unknown parameters without learning [2], adaptive decision rule [6], and sequential stochastic control S1 [7]. The estimator considered in this paper is ordinary least squares.

The two Monte Carlo experiments are presented in Section II. In the first experiment the terminal target is varied thus varying the value of anticipation. In the second experiment the dynamics of the model are varied. For both experiments the relative performance of the HCE and OLMV strategies varies between cases. The importance of active learning in the DUAL1 is also problem specific.

In the concluding section statistical inferences are drawn concerning the relative effectiveness of the alternative strategies. The fact that there is no dominance between the HCE and OLMV strategies is inferred. In comparing the OLMV and DUAL1, there is an indication that active learning can lead to poorer performance in problems where estimation error leads to overestimating the value of probing.

II. Single Equation Experiments

To gain insights into the performance of alternative estimation and control strategies, it is desirable to design experiments where a single parameter is varied. The first experiment was designed to investigate the effect of varying the value of the terminal state track. The effect of varying the terminal state track is to vary the importance of accurately estimating the true control law for the final decision, i.e. vary the importance of anticipating future observations on prior decisions. The second experiment was designed to investigate the effect of varying the dynamics of the system.

The specification of the two experiments is as follows:

Objective function
 Weights: $W_1 = 1.0$; $W_2 = 0.001$
 Time horizon: 10 periods
 Tracks: $\alpha_t = 0 \quad t = 1, 2, \ldots 9 \quad \alpha_{10}$ is defined below
 $\beta_t = -1.0 \quad t = 1, 2, \ldots 10$

Unknown system
 Equation: $X_t = \gamma_1 X_{t-1} + \gamma_2 U_t + \gamma_3 Z_t + \varepsilon_t$
 Disturbance: $\varepsilon_t \sim N(0, 1)$
 Exogenous variable: $Z_t = 1.0$ all t

Prior observations
 Number of prior observations: 5
 Initial state: $X_{-5} = 0$
 Sequence of fixed controls: $U_t = -1, -2, 0, 0, -2$. $t = -4, -3, \ldots, 0$
 Exogenous $Z_t = 1.0$ all t

Experiment 1: $\gamma_1 = 0.00001$, $\gamma_2 = 0.1$, $\gamma_3 = 1.0$

	Case 1	*Case 2*	*Case 3*	*Case 4*
α_{10}	0.0	10.0	100.0	1000.0

Experiment 2: $\alpha_{10} = 10.0$, $\gamma_1 = \gamma_2 = \gamma$, $\gamma_3 = 1.0$

	Case 1	*Case 2*	*Case 3*
γ	0.1	1.0	2.0

Examining the specifications of the experiments, the following items are noted: $W_2 \ll W_1$ which implies "cheap" control; tracks imply do-nothing until the final period; the system in experiment 1 has very little dynamics; and the prior observations were designed so that the variances of the parameters at the first decision were large.

Both experiments were evaluated by a Monte Carlo experiment for the three alternative estimation and control strategies together with the true stochastic control law which could be employed if the parameters were known, KNOWN. The KNOWN strategy provides a lower bound to judge the performance of the other strategies. A normal random number generator approximating $N(0, 1)$ was employed to generate the disturbances. The process started with generating prior observations so that each realization started at a different position with different initial estimates and covariances. For each case the Monte Carlo experiment is based on 100 realizations. The results for experiment 1 are as follows:

EXPERIMENT 1 RESULTS

	Case 1 $\alpha_{10}=0$		Case 2 $\alpha_{10}=10.0$		Case 3 $\alpha_{10}=100.0$		Case 4 $\alpha_{10}=1000.0$	
	Mean	Std Mean	Mean	Std Mean	Mean	Std Mean	Mean	Std Mean
KNOWN	5.25	0.2	5.65	0.2	54	0.2	4,947	1
HCE	17.36	3.8	31.98	4.0	1,726	226.7	161,511	22,096
OLMV	9.67	0.4	46.53	1.6	3,512	256.1	83,697	15,894
DUAL1	10.49	0.5	18.72	0.9	163	5.6	5,963	46

The relative performance of the DUAL1 to the HCE and OLMV strategies demonstrates the increasing importance of anticipation as α_{10} increases from 0 to 1000. With increasing values of α_{10} the last term dominates the objective function; hence it is not surprising that the performance of the DUAL1 appears to be converging towards the KNOWN as α_{10} increases.

The relative performance of the HCE and OLMV strategy can be attributed to the fact that these two strategies have very different learning characteristics. To discuss learning for problems involving more than one unknown parameter requires a learning statistic. One possibility is the F statistic[2] for the hypothesis that A and B are equal to zero. If we examine the case for $\alpha_{10}=100$ for 70 realizations, the HCE had a higher F statistic prior to the 10th decision and better performance. For only 7 realizations did the OLMV have both a higher F statistic prior to the 10th decision and better performance. For the case where $\alpha_{10}=100$ for 57 realizations, the OLMV had higher 10th period and better performance, whereas the same was true for the HCE in only 21 cases. As the OLMV strategy has covariances in both the numerator and denominator, it cannot be a priori assumed that the OLMV strategy is more "conservative" than the HCE strategy. As α_{10} is increased the OLMV strategy becomes less "conservative" that the HCE strategy especially in the 9th period decision. For $\alpha_{10}=10{,}000$ the OLMV strategy was superior to the HCE strategy.

The results for experiment 2 are as follows:

[2] This statistic may not be optimal as the relationship between the F statistic and performance is not known.

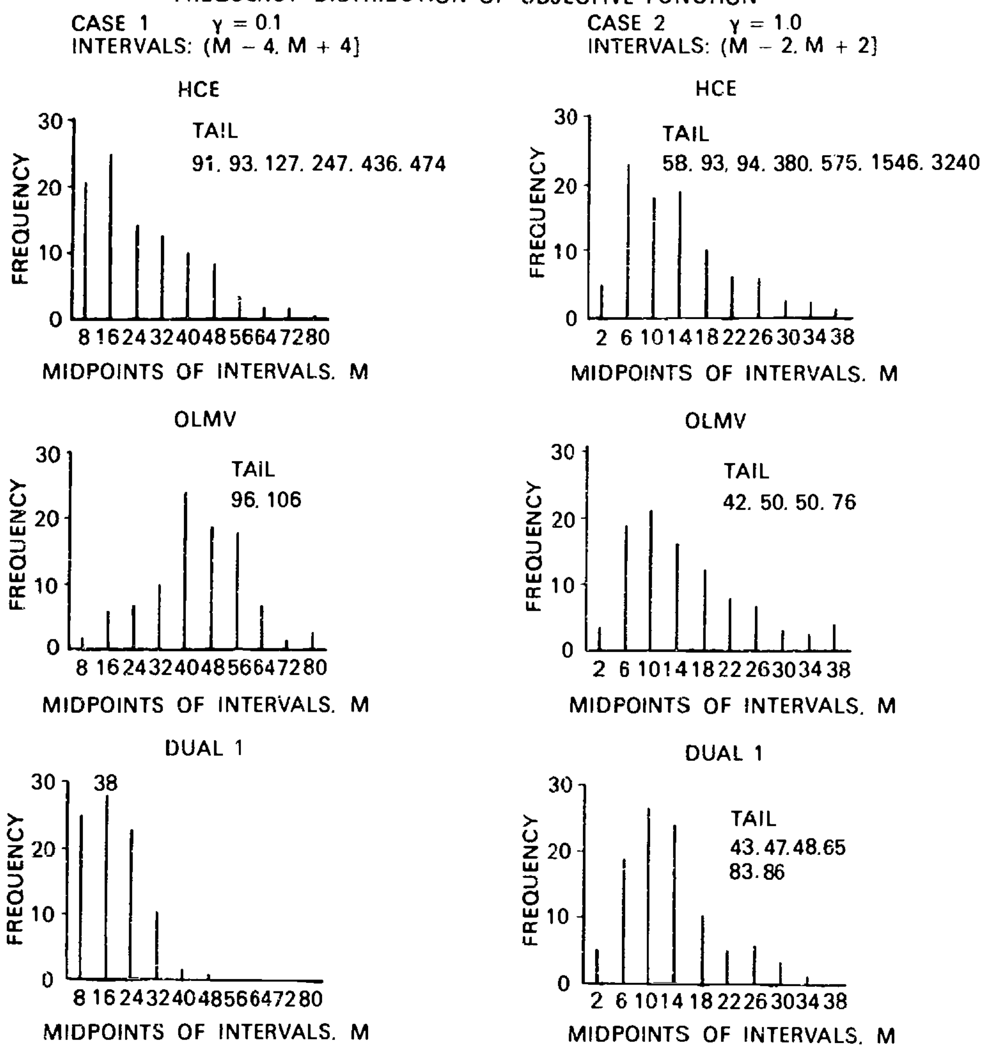
EXPERIMENT 2 RESULTS

	Case 1, $\gamma = .1$		Case 2, $\gamma = 1.0$		Case 3, $\gamma = 2.0$	
	Mean	Std Mean	Mean	Std Mean	Mean	Std Mean
KNOWN	5.65	0.2	5.22	0.2	5.3	0.2
HCE	38.45	6.7	72.1	36.0	454.8	153.0
OLMV	45.61	1.8	15.1	1.0	289.0	102.0
DUAL1	18.76	0.9	19.0	2.1	291.6	102.5

What is interesting about experiment 2 is that for case 2 the OLMV strategy outperforms the DUAL1 strategy. As an aid to discussion a frequency graph of the outcomes for case 1 and case 2 is displayed on figure 1.

EXPERIMENT 2

FREQUENCY DISTRIBUTION OF OBJECTIVE FUNCTION



As is explained in Appendix 1 the DUAL1 control algorithm weighs between two opposing methods of reducing the uncertainty: caution to reduce the effect of the path on present uncertainty and probing to reduce the effect of future uncertainty in the parameters. With increasing dynamics two effects will increase the importance of caution. First, increasing learning will generally take place without probing, thus decreasing the marginal value of probing. Second, probing will have a larger effect on the subsequent path, thus increasing the marginal cost.

Consider first the case where $\gamma = 0.1$. The median of the HCE strategy lies to the left of the OLMV strategy. The HCE strategy for this case is generally more active than the OLMV. With little dynamics the cost of probing is primarily one period. The HCE strategy generally outperforms the OLMV in the terminal period. The DUAL1 strategy appears to be probing in comparison to the OLMV strategy. With respect to the HCE strategy the DUAL1 frequently is more cautious.

What is interesting about case 2 is the performance of the OLMV strategy relative to the DUAL1 strategy. The distribution of the OLMV strategy appears to lie slightly to the left of the DUAL1 strategy distribution. An examination of the output reveals that for cases where the initial estimate of $\hat{B}$ is close to zero, the DUAL1 strategy frequently overestimates the value of probing. An example is shown below:

Example of excessive probing

	HCE	OLMV	DUAL1
$\hat{A}_0$	0.607	0.607	0.607
$\hat{B}_0$	0.157	0.157	0.157
U_1	0.631	−0.978	−8.109
X_1	0.304	−1.305	−8.437
OBJ	39.0	12.4	47.7

With $\hat{B}_0 = 0.157$ the DUAL1 seriously underestimates the impact probing on the subsequent path. In the second period the DUAL1 must correct the first period control, which has incurred a large cost on the first period state. From experiment 1 one would assume that α_{10} were increased from 10 to 1000, the performance of the DUAL1 strategy would improve relative to the OLMV strategy. The results for 20 realizations are as follows:

Experiment 2 Case 2a

$\gamma = 1.0 \quad \alpha_{10} = 1000$

	Mean	Std Mean
OLMV	10,225	5,153
DUAL1	2,737	601

Increasing α_{10} to 1000 greatly increases the value of anticipating the future observation pattern. Errors in the first period decision are dwarfed by the gain in performance in the final period.

The effect of increasing the dynamics from 0.1 to 1.0 on the relative performance of the HCE strategy to the OLMV strategy is twofold. The increased dynamics generate more extreme values in the tail of the HCE distribution and at the same time increase the passive learning of the OLMV strategy to a more nearly optimal level.

In case 3 the DUAL1 strategy utilizes very little probing. An example follows:

	HCE	OLMV	DUAL1
U_1	72.06	65.151	65.161
OBJ	1520.0	855.7	856.5

With $\gamma = 2.0$ the DUAL1 strategy probes slightly more than the OLMV strategy. If γ is increased to 3.0, the DUAL1 strategy is slightly more cautious than the OLMV strategy.

III. Conclusions

For the Monte Carlo experiments statistical inference can be made concerning the relative merits of the alternative estimation and control strategies. In describing the tests the expression DUAL1 > OLMV means

$$H_0: \text{Mean}_{\text{DUAL1}} \geq \text{Mean}_{\text{OLMV}} \text{ versus } H_1: \text{Mean}_{\text{DUAL1}} < \text{Mean}_{\text{OLMV}}.$$

The test employed is a t-test of the difference of the two means for paired observations. To test whether the performance of the HCE and OLMV strategy is problem specific the following tests were considered

Experiment	Test	t-statistic
1 Case 4	OLMV > HCE	−2.8
1 Case 2	HCE > OLMV	−3.6

The conclusion is reached that the performance of the HCE and OLMV is problem specific. This result amplifies the previous Monte Carlo experiments in HCE and OLMV strategies [6], [7]. In [6] Prescott found that the OLMV strategy was superior to the HCE for each problem considered. In [7] Sarris and Athans have an example with constant coefficients where the mean of HCE strategy is lower than the mean of the OLMV for 20 realizations. In comparison with Monte Carlo experiments of other passive learning strategies [3], [8] the conjecture is reached that it is unlikely a particular passive learning strategy will dominate its competitors.

The results of Experiment 2 Case 2 raise the issue of whether a passive learning strategy can produce better results than an active learning strategy. The t statistic for OLMV > DUAL1 is −2.0 Experiment 2 Case 2 was repeated twice with the following results:

	OLMV		DUAL1		
Case	Mean	Std	Mean	Std	t for OLMV > DUAL1
2b	16.15	1.3	22.85	3.7	−1.91
2c	16.66	1.6	19.83	1.7	−2.09
Sum	15.94	0.8	20.54	1.5	−3.22

From these results it can be assumed the OLMV strategy is superior to the DUAL1 strategy for this problem.

University of Texas at Austin

References

[1] Bar-Shalom, Y. and E. Tse. "Concepts and Methods in Stochastic Control," a chapter in *Control and Dynamic Systems: Advances in Theory and Applications*, C. T. Leondes, (ed), Academic Press, 1975.
[2] Chow, G. C., *Analysis and Control of Dynamic Economic Systems*, New York: Wiley, 1975.
[3] Ku, R. and M. Athans, "On the Adaptive Control of Linear Systems Using the Open-Loop-Feedback Optimal Approach," *IEEE Trans. on Automatic Control*, Vol. AC-18, No. 5, 1975.
[4] MacRae, Elizabeth Chase, "Linear Decision with Experimentation," *Annals of Economic and Social Measurement*, 1:4, 1972.
[5] Norman, A. L. and W. S. Jung, "Linear Quadratic Control for Models with Long Lags," forthcoming *Econometrica*.
[6] Prescott, E., "Adaptive Decision Rules for Macroeconomic Planning," *Western Economic Review*, 1971, Vol. 9, p. 369–378.
[7] Sarris, A. H. and M. Athans, "Optimal Adaptive Control Methods for Structurally Varying Systems." Working Paper #24, National Bureau of Economic Research, Cambridge, Mass., Dec. 1973.
[8] Tse, Edison and M. Athans, "Adaptive Stochastic Control for a Class of Linear Systems," *IEEE Trans. on Automatic Control*, Vol. AC-17, No. 1, Feb. 1972.
[9] Tse, Edison and Y. Bar-Shalom, "An Actively Adaptive Control for Linear Systems with Random Parameters via the Dual Control Approach," *IEEE Trans. on Automatic Control*, Vol. AC-18, No. 2, April 1973.

Appendix 1

Like the Tse, Bar-Shalom Dual Control, the DUAL1 approach divides the problem to be solved into the current control, future deterministic control, and future perturbation control.

The assumed dynamics for the DUAL1 are

(A.1) $$X_t = \tilde{A}_t X_{t-1} + \tilde{B}_t U_t + \tilde{C}_t Z_t + \varepsilon_t$$

$\tilde{A}_t, \tilde{B}_t, \tilde{C}_t$ are random variables whose means are equal to the estimates obtained from the observed data and whose covariances are obtained from the observed data plus future nominal data through $t-1$. Partitioning the random variables into their deterministic and random components, e.g.,

$$X_t = \overline{X}_t + \Delta X_t, \quad \tilde{A}_t = \overline{A} + \Delta A_t, \text{ etc.},$$

the deterministic component can be written

(A.2) $$\overline{X}_t = \overline{AX}_{t-1} + \overline{BU}_t + \overline{C}Z_t$$

and ignoring second order terms, e.g., $\Delta A_t \Delta X_t$, the stochastic component is

(A.3) $$\Delta X_t = \overline{A}\Delta X_{t-1} + \overline{B}\Delta U_t + \Delta A_t \overline{X}_{t-1} + \Delta B_t \overline{U}_t + \Delta C_t Z + \varepsilon_t.$$

The deterministic component problem is

(A.4) $$\text{Determine } J^*(\overline{X}_k) = \min_{\bar{U}_{k+1}, \ldots, \bar{U}_t} J(\overline{X}_k)$$

where

$$J = \sum_{t=k+1}^{T} [\tfrac{1}{2}(\overline{X}_t - \alpha_t)' W_1 (\overline{X}_t - \alpha_t) + \tfrac{1}{2}(\overline{U}_t - \beta_t)' W_2 (\overline{U}_t - \beta_t)] \tag{A.5}$$

subject to

$$\overline{X}_t = \overline{AX}_{t-1} + \overline{BU}_t + \overline{C}Z_t. \tag{A.6}$$

By a straightforward application of recursive dynamic programming the solution of this problem can be written as

$$J^*(\overline{X}_k) = Q_1(k+1) + \overline{X}_k' Q_2(k+1) + \tfrac{1}{2}\overline{X}_k' Q_3(k+1)\overline{X}_k. \tag{A.7}$$

The stochastic component problem is

$$J^*(\Delta X_k) = \min_{\Delta U_{k+1},\ldots,\Delta U_t} J(\Delta X_k) \tag{A.8}$$

where

$$J = E\left\{\sum_{t=k+1}^{T} (\tfrac{1}{2}\Delta X_t' W_1 \Delta X_t + \Delta X_t' W_1 (\overline{X}_t - \alpha_t) + \tfrac{1}{2}\Delta U_t' W_2 \Delta U_t + \Delta U_t' W_2 (\overline{U}_t - \beta_t) | \mathcal{P}_{t-1})\right\} \tag{A.9}$$

subject to

$$\Delta X_t = \overline{A}\Delta X_{t-1} + \overline{B}\Delta U_t + \Delta A_t \overline{X}_{t-1} + \Delta B_t \overline{U}_t + \Delta C_t Z_t + \varepsilon_t. \tag{A.10}$$

Proceeding by the usual recursive dynamic programming formulation assume the solution can be expressed as a quadratic form

$$J^*(\Delta X_j) = Q_4(j+1) + \Delta X_j' Q_5(j+1) + \Delta X_j' Q_6(j+1)\overline{X}_j + \tfrac{1}{2}\Delta X_j' Q_7(j+1)\Delta X_j \tag{A.11}$$

then,

$$\begin{aligned} J^*(\Delta X_{j-1}) = \min_{\Delta U_j} E\{&\tfrac{1}{2}(\Delta X_j' W_1 \Delta X_j + \Delta X_j' W_1 (\overline{X}_j - \alpha_j) + \tfrac{1}{2}\Delta U_j' W_2 \Delta U_j \\ &+ \Delta U_j' W_2 (\overline{U}_j - \beta_j) + Q_4(j+1) + \Delta X_j' Q_5(j+1) \\ &+ \Delta X_j' Q_6(j+1)\overline{X}_j + \tfrac{1}{2}\Delta X_j' Q_7(j+1)\Delta X_j | \mathcal{P}_{j-1}\}. \end{aligned} \tag{A.12}$$

Substitution for ΔX_j and collecting terms

$$\begin{aligned} J^*(\Delta X_{j-1}) = \min_{\Delta U_j} [&\tfrac{1}{2}(\overline{A}\Delta X_{j-1} + \overline{B}\Delta U_j)'(W_1 + Q_7(j+1))(\overline{A}\Delta X_{j-1} + \overline{B}\Delta U_j) \\ &+ (\overline{A}\Delta X_{j-1} + \overline{B}\Delta U_j)'[(W_1 + Q_6(j+1))\overline{X}_j - W_1\alpha_j + Q_5(j+1)] \\ &+ \tfrac{1}{2}\Delta U_j' W_2 \Delta U_j + \Delta U_j' W_2 (\overline{U}_j - \beta_j) \\ &+ \Phi_j + E\{\varepsilon_j'(W_1 + Q_7(j+1))\varepsilon_j\} + Q_4(j+1)] \end{aligned} \tag{A.13}$$

where

$$\Phi_j = E\{\tfrac{1}{2}(\Delta A_j\bar{X}_{j-1} + \Delta B_j\bar{U}_j + \Delta C_j Z_j)'(W_1 + Q_7(j+1))(\Delta A_j\bar{X}_{j-1} + \Delta B_j\bar{U}_j + \Delta C_j Z_j).$$

Solving for ΔU_j

$$\Delta U_j = -[\bar{B}'(W_1 + Q_7(j+1))\bar{B} + W_2]^{-1}[\bar{B}'(W_1 + Q_7(j+1))\bar{A}\Delta X_{j-1} + \bar{B}'\{W_1 + Q_6(j+1))\bar{X}_j - W_1\alpha_j + Q_5(j+1)\} + W_2(\bar{U}_j - \beta_j)]. \tag{A.14}$$

Let

$$S_{j+1} = W_1 + Q_7(j+1); \quad D_j = [\bar{B}'S_{j+1}\bar{B} + W_2]; \quad G_j = D_j^{-1}[\bar{B}'S_{j+1}\bar{A}\Delta X_{j-1}] \tag{A.15}$$

Assuming $Q_7(j+1) = Q_6(j+1)$

$Q_6(j+1) = Q_3(j+1)$ (from the deterministic component)

$Q_5(j+1) = Q_2(j+1)$ (from the deterministic component)

then

$$\Delta U_j = G_j\Delta X_{j-1} - D_j^{-1}[B'(S_{j+1})\bar{X}_j - W_1\alpha_j + Q_2(j+1) + W_2(\bar{U}_j - \beta_j)]. \tag{A.16}$$

Substituting for $\bar{X}_j$ and $\bar{U}_j$, (A.16) can be reduced to

$$\Delta U_j = G_j\Delta X_{j-1} \tag{A.17}$$

where G_j is the same as the deterministic component.

Substituting for ΔU_j, $\bar{X}_j$ and $\bar{U}_j$ in (A.13) and collecting terms

$$Q_7(j) = \bar{A}'S_{j+1}\bar{A} + G_j'D_jG_j = Q_3(j) \tag{A.18}$$

$$Q_6(j) = \bar{A}'S_{j+1}\bar{A} + G_j'D_jG_j = Q_3(j) \quad \text{from the deterministic component} \tag{A.19}$$

$$\begin{aligned} Q_5(j) &= G_j'B'S_{j+1}\bar{B}g_j + \bar{A}'S_{j+1}\bar{B}g_j + G_j'\bar{B}'S_{j+1}\bar{C}Z_j + \bar{A}'S_{j+1}\bar{C}Z_j \\ &\quad + \bar{A}'(Q_2(j+1) - W_j\alpha_j) + G_j'\bar{B}_j'(Q_2(j+1) - W_1\alpha_j) + G_j'W_2(g_j - \beta_j) \\ &= Q_2(j) \quad \text{(from the deterministic component)} \end{aligned} \tag{A.20}$$

$$Q_4(j) = Q_4(j+1) + \Phi_j + E\{\varepsilon_j'S_{j+1}\varepsilon_j\}. \tag{A.21}$$

Thus, the cost to go, i.e., perturbation plus deterministic component problems can be written

$$\begin{aligned} J^c(X_k) &= J^*(\bar{X}_k) + J^*(\Delta X_k) \\ &= Q_1(k+1) + Q_4(k+1) + X_k'Q_2(k+1) + \tfrac{1}{2}X_k'Q_3X_k. \end{aligned} \tag{A.22}$$

The only term which depends on $\bar{X}_t$ and $\bar{U}_t$ is $Q_4(k+1)$. Given the solution to the perturbation control problem, U_k is obtained from

$$\begin{aligned} J^{*D}(X_{k-1}) = \min_{U_k} E\{&\tfrac{1}{2}(X_k - \alpha_k)'W_1(X_k - \alpha_k) + \tfrac{1}{2}(U_k - \beta_k)'W_2(U_k - \beta_k) \\ &+ J^C(X_k)|\mathcal{P}_{k-1}\} \end{aligned} \tag{A.23}$$

$$= \min_{U_k} [\tfrac{1}{2}(\overline{X}_k - \alpha_k)' W_1(\overline{X}_k - \alpha_k) + \tfrac{1}{2}(U_k - \beta_k)' W_2(U_k - \beta_k) + Q_4(k)$$

$$+ Q_1(k+1) + \overline{X}_k Q_2(k+1) + \overline{X}_k Q_3(k+1)\overline{X}_k].$$

Because Φ_j's must be evaluated along the future nominal path, U_k must be obtained by numerical methods. For the first order DUAL1, $Q_2(k+1)$ and $Q_3(k+1)$ are equivalent for the deterministic and perturbation component problems; hence, $Q_1(k+1)$, $Q_2(k+1)$ and $Q_3(k+1)$ need be solved only once to compute U_k. For the example shown, a quadratic fit linear search was employed to compute single control variable problems and a quasi-Newton algorithm can be employed for multi-variable problems. It should be noted that there are two considerations in reducing the value of Φ_j: the nominal path $[\overline{X}_{j-1}, \overline{U}_j, \overline{Z}_j]$ and the covariance elements of the $[\Delta A_j, \Delta B_j, \Delta C_j]$.

To compare the DUAL1 with the MacRae adaptive control Table 2 of [4] was computed for the DUAL1 with the following results:

TABLE 2 of [7]

FIRST PERIOD POLICIES FOR DIFFERENT HORIZON LENGTHS

N = Horizon; Goals = 0

$a = 0.7$ $b = -0.5$ $c = 3.5$
$\Omega = 0.2$ $\Gamma_0^{bb} = 0.5$ $X_0 = 0.0$

	$N=2$		$N=4$		$N=8$		$N=16$	
$q:r$	DUAL1	Adapt	DUAL1	Adapt	DUAL1	Adapt	DUAL1	Adapt
1:5	0.622	0.622	1.091	1.082	1.442	1.394	1.547	1.460
5:5	1.747	1.740	2.521	2.449	2.920	2.688	3.096	2.705
5:1	2.707	2.682	3.245	3.056	3.456	3.083	3.688	3.084
5:0	3.206	3.138	3.351	3.146	3.531	3.147	3.725	3.147

The DUAL1 control is slightly more "aggressive" than the MacRae adaptive control. Bar-Shalom and Tse have examined the case $q:r$ is 5:5 and $N = 2$ in a Monte Carlo experiment. They show that the original version of the dual control produces a first period decision of 1.33. The basic difference between DUAL1 and the Tse and Bar-Shalom dual is the fact that the DUAL1 contains no covariances of the state and unknown parameters. If this term is eliminated from the objective function for the Tse and Bar-Shalom dual control, the first period decision is 1.746.